

50C

RAW SEQUENCE LISTING

The Biotechnology Systems Branch of the Scientific and Technical Information Center (STIC) no errors detected.

Application Serial Number: 10/522,389
Source: PG
Date Processed by STIC: 1/26/06

ENTERED



PCT

RAW SEQUENCE LISTING

DATE: 01/26/2006

PATENT APPLICATION: US/10/522,389

TIME: 09:19:40

Input Set : F:\24615-20249.00 - Seqlist (client).txt

Output Set: N:\CRF4\01262006\J522389.raw

3 <110> APPLICANT: PEL, Herman J.
 4 HOPPER, Sylvia
 6 <120> TITLE OF INVENTION: GENES FROM PROPIONIBACTERIUM FREUDENREICHII
 7 ENCODING ENZYMES INVOLVED IN VITAMIN B12 BIOSYNTHESIS
 9 <130> FILE REFERENCE: 246152024900
 11 <140> CURRENT APPLICATION NUMBER: US 10/522,389
 C--> 12 <141> **CURRENT FILING DATE: 2005-01-25**
 14 <150> PRIOR APPLICATION NUMBER: PCT/EP03/008216
 15 <151> PRIOR FILING DATE: 2003-07-25
 17 <150> PRIOR APPLICATION NUMBER: EP 02255203.8
 18 <151> PRIOR FILING DATE: 2002-07-25
 20 <160> NUMBER OF SEQ ID NOS: 17
 22 <170> SOFTWARE: PatentIn version 3.1
 24 <210> SEQ ID NO: 1
 25 <211> LENGTH: 2586
 26 <212> TYPE: DNA
 27 <213> ORGANISM: Propionibacterium freudenreichii
 29 <220> FEATURE:
 30 <221> NAME/KEY: CDS
 31 <222> LOCATION: (1)..(2586)
 32 <223> OTHER INFORMATION:
 W--> 34 <400> 1
 35 atg gtg acg gcg acg gct ctt ccg cgg gtg ctc atc gcg gcc ccc gcg 48
 36 Met Val Thr Ala Thr Ala Leu Pro Arg Val Leu Ile Ala Ala Pro Ala
 37 1 5 10 15
 39 tcc agc cag gga aag acc acc gtg gcc atc ggc ctg atg gcg gcc ctg 96
 40 Ser Ser Gln Gly Lys Thr Thr Val Ala Ile Gly Leu Met Ala Ala Leu
 41 20 25 30
 43 cgg gcc tcg ggg cgc agc gtg gcc gga ttc aag gtg ggc ccc gac tac 144
 44 Arg Ala Ser Gly Arg Ser Val Ala Gly Phe Lys Val Gly Pro Asp Tyr
 45 35 40 45
 47 atc gat ccg ggc tat cac gca ctg gcc tgc ggt cgc ccc ggc cgc aac 192
 48 Ile Asp Pro Gly Tyr His Ala Leu Ala Cys Gly Arg Pro Gly Arg Asn
 49 50 55 60
 51 ctg gat ccc tat ttg tgc ggg ccc gag cgc att gcg ccg ttg ttc gcc 240
 52 Leu Asp Pro Tyr Leu Cys Gly Pro Glu Arg Ile Ala Pro Leu Phe Ala
 53 65 70 75 80
 55 cat ggc gcg ctg cat ccc gaa ccc gcg gac atc tcg gtc gtc gaa ggc 288
 56 His Gly Ala Leu His Pro Glu Pro Ala Asp Ile Ser Val Val Glu Gly
 57 85 90 95
 59 gtg atg ggc atg ttc gac ggc aag ctc ggc gcg tgg ccc gac ggc acc 336
 60 Val Met Gly Met Phe Asp Gly Lys Leu Gly Ala Trp Pro Asp Gly Thr
 61 100 105 110

RAW SEQUENCE LISTING

PATENT APPLICATION: US/10/522,389

DATE: 01/26/2006

TIME: 09:19:40

Input Set : F:\24615-20249.00 - Seqlist (client).txt

Output Set: N:\CRF4\01262006\J522389.raw

```

63 gat gac ccc gcc ggt ttt ggc tca tgc gcc cat atc gcc agg ctg ctc      384
64 Asp Asp Pro Ala Gly Phe Gly Ser Ser Ala His Ile Ala Arg Leu Leu
65      115      120      125
67 gat gcc ccc gtg ctg ctc gtg gtc gac ggc tca cac agt gcc cgt acc      432
68 Asp Ala Pro Val Leu Leu Val Val Asp Gly Ser His Ser Ala Arg Thr
69      130      135      140
71 gcc gca gcc ctg tgc cat ggc ctg gcc agc tac gat ccc cgc atc cat      480
72 Ala Ala Ala Leu Cys His Gly Leu Ala Ser Tyr Asp Pro Arg Ile His
73 145      150      155      160
75 gtg gcc ggc gtc atc ctc aat cgg gtg atg ggt gcc cgc gtg gtc gac      528
76 Val Ala Gly Val Ile Leu Asn Arg Val Met Gly Ala Arg Val Val Asp
77      165      170      175
79 gag atc acc cgg ggc tgc gca cgt gtc ggc ctg ccg gtg ctg ggg gct      576
80 Glu Ile Thr Arg Gly Cys Ala Arg Val Gly Leu Pro Val Leu Gly Ala
81      180      185      190
83 ctg ccg aaa agc acg cgg gtg gcc gtg ggc tca cgc cac ctg gga ctg      624
84 Leu Pro Lys Ser Thr Arg Val Ala Val Gly Ser Arg His Leu Gly Leu
85      195      200      205
87 gtc acg gcc gac gag cag ggt gac gcg atc ggc atc gtg cag cag gcc      672
88 Val Thr Ala Asp Glu Gln Gly Asp Ala Ile Gly Ile Val Gln Gln Ala
89      210      215      220
91 ggt gag ctc gtc gcc gca cac ctc gac ctc gac gcc atc gcc acg atc      720
92 Gly Glu Leu Val Ala Ala His Leu Asp Leu Asp Ala Ile Ala Thr Ile
93 225      230      235      240
95 gcc ggt ggg gcc cct gac ctg gcc gtc gat ccc tgg gat ccc gcc gca      768
96 Ala Gly Gly Ala Pro Asp Leu Ala Val Asp Pro Trp Asp Pro Ala Ala
97      245      250      255
99 gag gtc gaa ccg gta ccg ggg cgt ccg gtc atc gcc atg gcc tgc ggt      816
100 Glu Val Glu Pro Val Pro Gly Arg Pro Val Ile Ala Met Ala Ser Gly
101      260      265      270
103 ccc gca ttc acc ttc cgg tac acc gaa acc gca gaa ctg ctg gag gcg      864
104 Pro Ala Phe Thr Phe Arg Tyr Thr Glu Thr Ala Glu Leu Leu Glu Ala
105      275      280      285
107 gcc ggc tgc cgg gtg acg gcc ttc gat ccg ctc acc gcc cgg ggc ctt      912
108 Ala Gly Cys Arg Val Thr Ala Phe Asp Pro Leu Thr Ala Arg Gly Leu
109      290      295      300
111 ccg gcc gat gtg tcc ggc ctg tac ctg ggg ggt ggt ttc ccc gag gag      960
112 Pro Ala Asp Val Ser Gly Leu Tyr Leu Gly Gly Gly Phe Pro Glu Glu
113 305      310      315      320
115 cac gcc gag gcg ctc gcc ggc aac acc tcc ctg ggc gct gaa atc gcc      1008
116 His Ala Glu Ala Leu Ala Gly Asn Thr Ser Leu Gly Ala Glu Ile Ala
117      325      330      335
119 tca cgc gtg tcc gag ggc ctg ccg acg gtg gcc gag tgt gcg ggg ctg      1056
120 Ser Arg Val Ser Glu Gly Leu Pro Thr Val Ala Glu Cys Ala Gly Leu
121      340      345      350
123 ctc tac ctg tgc cgc agc ctg gat gga ctg gcg atg gcc ggg gtg gtc      1104
124 Leu Tyr Leu Cys Arg Ser Leu Asp Gly Leu Ala Met Ala Gly Val Val
125      355      360      365
127 gac gcc gac tgc tcc atg acg ccg cgc ctg acc atc ggc tac cac cac      1152

```

RAW SEQUENCE LISTING

PATENT APPLICATION: US/10/522,389

DATE: 01/26/2006

TIME: 09:19:40

Input Set : F:\24615-20249.00 - Seqlist (client).txt

Output Set: N:\CRF4\01262006\J522389.raw

128	Asp	Ala	Asp	Ser	Ser	Met	Thr	Pro	Arg	Leu	Thr	Ile	Gly	Tyr	His	His	
129		370					375					380					
131	gcc	cgc	gcc	gcc	aac	gac	agc	ttc	ctg	atg	cgc	gcc	ggg	gag	cgc	tat	1200
132	Ala	Arg	Ala	Ala	Asn	Asp	Ser	Phe	Leu	Met	Arg	Ala	Gly	Glu	Arg	Tyr	
133	385				390						395					400	
135	cgg	gcc	cat	gag	ttc	cac	cgc	acc	acc	ctg	gac	acg	ccc	ccc	tac	gac	1248
136	Arg	Ala	His	Glu	Phe	His	Arg	Thr	Thr	Leu	Asp	Thr	Pro	Pro	Tyr	Asp	
137				405						410					415		
139	cgc	gac	ccc	gga	cca	caa	cgg	ctg	ggc	gac	caa	cgg	ttg	gcg	tgg	gac	1296
140	Arg	Asp	Pro	Gly	Pro	Gln	Arg	Leu	Gly	Asp	Gln	Arg	Leu	Ala	Trp	Asp	
141			420						425					430			
143	gtg	gag	acc	ccg	acg	ggg	ggc	aac	cga	ccc	gag	ggg	gtg	ctg	gtc	gcc	1344
144	Val	Glu	Thr	Pro	Thr	Gly	Gly	Asn	Arg	Pro	Glu	Gly	Val	Leu	Val	Ala	
145		435					440					445					
147	ccg	acc	ccc	ggt	tcc	gcg	ccc	agc	gtc	cac	gcc	tcc	tac	cag	cac	ctg	1392
148	Pro	Thr	Pro	Gly	Ser	Ala	Pro	Ser	Val	His	Ala	Ser	Tyr	Gln	His	Leu	
149	450					455					460						
151	cac	tgg	gca	ggg	agt	ccg	gta	ctg	gcg	caa	cgc	ttc	gcc	cgg	gcg	gcg	1440
152	His	Trp	Ala	Gly	Ser	Pro	Val	Leu	Ala	Gln	Arg	Phe	Ala	Arg	Ala	Ala	
153	465				470					475					480		
155	agc	gaa	tat	ggg	cac	acc	ggc	cat	cac	tcc	ccc	cgg	cct	gcc	gcc	acg	1488
156	Ser	Glu	Tyr	Gly	His	Thr	Gly	His	His	Ser	Pro	Arg	Pro	Ala	Ala	Thr	
157			485						490					495			
159	acg	ccg	gga	gat	gcg	ttg	tcc	gca	gcg	ccc	gac	ctc	acc	cat	cac	ggg	1536
160	Thr	Pro	Gly	Asp	Ala	Leu	Ser	Ala	Ala	Pro	Asp	Leu	Thr	His	His	Gly	
161			500						505					510			
163	gat	cgc	gat	gtg	ctg	ccc	ggc	ctg	gtc	gac	ttg	gcg	gtg	aac	gtg	cgc	1584
164	Asp	Arg	Asp	Val	Leu	Pro	Gly	Leu	Val	Asp	Leu	Ala	Val	Asn	Val	Arg	
165		515						520					525				
167	gat	gtg	cga	cct	ccg	gcc	tgg	ctc	gtg	gag	cgc	atc	gtc	gcc	tcc	agc	1632
168	Asp	Val	Arg	Pro	Pro	Ala	Trp	Leu	Val	Glu	Arg	Ile	Val	Ala	Ser	Ser	
169	530					535					540						
171	gac	cag	tgg	gcc	cac	tac	ccc	gat	cag	cgc	gaa	gcg	acc	cgt	gcg	gtg	1680
172	Asp	Gln	Trp	Ala	His	Tyr	Pro	Asp	Gln	Arg	Glu	Ala	Thr	Arg	Ala	Val	
173	545				550					555					560		
175	gca	ctg	cgc	cat	ggc	gtc	aac	ccc	gac	cag	gta	ctg	ctc	acg	gcc	ggg	1728
176	Ala	Leu	Arg	His	Gly	Val	Asn	Pro	Asp	Gln	Val	Leu	Leu	Thr	Ala	Gly	
177			565						570					575			
179	tcc	tcg	gag	gcg	ttc	agc	ctg	atc	gcc	cac	ggg	ttc	tcc	ccg	cgc	tgg	1776
180	Ser	Ser	Glu	Ala	Phe	Ser	Leu	Ile	Ala	His	Gly	Phe	Ser	Pro	Arg	Trp	
181			580						585					590			
183	gcg	gtc	gtg	gtg	cat	ccc	cag	ttc	acc	gaa	cca	gag	gtg	gcc	ctg	cgc	1824
184	Ala	Val	Val	Val	His	Pro	Gln	Phe	Thr	Glu	Pro	Glu	Val	Ala	Leu	Arg	
185		595					600						605				
187	aac	gcc	ggg	cgc	ccg	gtc	ggc	cgc	ctg	gtg	ctc	cat	gcc	tcg	gat	ggc	1872
188	Asn	Ala	Gly	Arg	Pro	Val	Gly	Arg	Leu	Val	Leu	His	Ala	Ser	Asp	Gly	
189	610					615					620						
191	ttc	cag	ttc	gat	cac	gaa	ctg	ctg	gac	ccc	agg	gcc	gac	atg	gtg	gtc	1920
192	Phe	Gln	Phe	Asp	His	Glu	Leu	Leu	Asp	Pro	Arg	Ala	Asp	Met	Val	Val	

RAW SEQUENCE LISTING

PATENT APPLICATION: US/10/522,389

DATE: 01/26/2006

TIME: 09:19:40

Input Set : F:\24615-20249.00 - Seqlist (client).txt

Output Set: N:\CRF4\01262006\J522389.raw

```

193 625          630          635          640
195 atc ggc aat ccg acc aat ccc acc ggc gtg ctg cat tcg gcg gcg agc 1968
196 Ile Gly Asn Pro Thr Asn Pro Thr Gly Val Leu His Ser Ala Ala Ser
197          645          650          655
199 ctg cgc gcg ttg tgc cgg ccc gga cgc gtg gtg gtt gac gag gca 2016
200 Leu Arg Ala Leu Cys Arg Pro Gly Arg Val Val Val Val Asp Glu Ala
201          660          665          670
203 ttc atg gac gcc gtg ccg ggc gag ccc gag agc ctc atc ggg gca cgc 2064
204 Phe Met Asp Ala Val Pro Gly Glu Pro Glu Ser Leu Ile Gly Ala Arg
205          675          680          685
207 atg gat ggc ctg ttg gtc acc cgc tcg ttc acg aag act tgg agc gtc 2112
208 Met Asp Gly Leu Leu Val Thr Arg Ser Phe Thr Lys Thr Trp Ser Val
209          690          695          700
211 ccg ggg ctg cgg atc gga tat gtg gtc ggg gat ccc gcg ctc att cgc 2160
212 Pro Gly Leu Arg Ile Gly Tyr Val Val Gly Asp Pro Ala Leu Ile Arg
213 705          710          715          720
215 gtc ctg gcg cac gaa cag ccc tgt tgg ccc atc tcc acc ccc gcc ctg 2208
216 Val Leu Ala His Glu Gln Pro Cys Trp Pro Ile Ser Thr Pro Ala Leu
217          725          730          735
219 gtc acc gcc cgc gaa tgc tcc acg cca cgc gcc gtg gag cag gcc acc 2256
220 Val Thr Ala Arg Glu Cys Ser Thr Pro Arg Ala Val Glu Gln Ala Thr
221          740          745          750
223 tca gat gcc cga cag gcg gcg cag gac cgc cga cac ctg gtg gcc cgc 2304
224 Ser Asp Ala Arg Gln Ala Ala Gln Asp Arg Arg His Leu Val Ala Arg
225          755          760          765
227 ctg gcc ggg atc ggc atc cag acc gtc ggg gag gcc agg gcc ccc ttc 2352
228 Leu Ala Gly Ile Gly Ile Gln Thr Val Gly Glu Ala Arg Ala Pro Phe
229          770          775          780
231 gtc cta gtc gac ctg cgc gcc cac ccg ccc ggt ggg ctt cgt gcg gga 2400
232 Val Leu Val Asp Leu Arg Ala His Pro Pro Gly Gly Leu Arg Ala Gly
233 785          790          795          800
235 ttg cgg acg ctc ggc ttc acc gtg cgc agc ggc gag agc ttc ccc ggc 2448
236 Leu Arg Thr Leu Gly Phe Thr Val Arg Ser Gly Glu Ser Phe Pro Gly
237          805          810          815
239 ctg ggc gcg ggc tgg ttg cgg ctc gcg gtt cgc cac ccg gac atc agc 2496
240 Leu Gly Ala Gly Trp Leu Arg Leu Ala Val Arg His Pro Asp Ile Ser
241          820          825          830
243 gac gcg ttc gtc gct gcc ctg gcc cgc acc atc gac gca ctg gac aca 2544
244 Asp Ala Phe Val Ala Ala Leu Ala Arg Thr Ile Asp Ala Leu Asp Thr
245          835          840          845
247 gcg cag cac ccc atg cga cca cca caa gga gac atc aga tga 2586
248 Ala Gln His Pro Met Arg Pro Pro Gln Gly Asp Ile Arg
249          850          855          860
252 <210> SEQ ID NO: 2
253 <211> LENGTH: 861
254 <212> TYPE: PRT
255 <213> ORGANISM: Propionibacterium freudenreichii
257 <400> SEQUENCE: 2
258 Met Val Thr Ala Thr Ala Leu Pro Arg Val Leu Ile Ala Ala Pro Ala

```


VERIFICATION SUMMARY

DATE: 01/26/2006

PATENT APPLICATION: US/10/522,389

TIME: 09:19:42

Input Set : F:\24615-20249.00 - Seqlist (client).txt

Output Set: N:\CRF4\01262006\J522389.raw

L:12 M:271 C: Current Filing Date differs, Replaced Current Filing Date
L:34 M:258 W: Mandatory Feature missing, <223> Blank for SEQ#:1,Line#:32
L:431 M:258 W: Mandatory Feature missing, <223> Blank for SEQ#:3,Line#:429
L:548 M:258 W: Mandatory Feature missing, <223> Blank for SEQ#:5,Line#:546
L:685 M:258 W: Mandatory Feature missing, <223> Blank for SEQ#:7,Line#:683